

EDITORIAL

AI IN PSYCHIATRY

Do Likes Reinforce Depression?

Quentin J. M. Huys, MD, PhD; Amy Orben, MA, DPhil

Social interaction is a critical component of mental well-being. Over the past few years, engagement via a variety of virtual modes has added new components to traditional social activities, and reduced the barriers for individuals to stay in touch with others and reach comparatively large audiences. The rapid growth of social media suggests that



Related article

people interacting with these services might find them rewarding. Rewards, by definition, are outcomes which reinforce preceding activities and promote repetition. However, there has been a growing chorus of concerns around social media, suggesting they might be harmful for both mental health and the social fabric of societies more broadly.¹ Indeed, US court judgments have come to the view that aspects of social media were explicitly designed to induce addictive behaviors.²

Mirea and colleagues³ present evidence speaking to this issue in the context of depressive symptoms and likes on Twitter, now known as X. Likes are one of the key mechanisms to promote and maintain engagement in social media. They provide a way of expressing agreement or support for others' posts and serve as a signal of support or engagement by others. By supporting engagement, likes are likely to function as reinforcers,⁴ eg, for moral outrage.⁵ It is clear that likes represent a design feature that reinforces social media behaviors⁶ and maximizes the time users spend on these platforms.

The situation for depressive symptoms is, however, particularly interesting, as a core feature of depression is a reduction in the enjoyment derived from previously enjoyable things.⁷ As such, and as seen in numerous laboratory studies,⁸ individuals with depressive symptoms might a priori be expected to be less sensitive to likes, which may even have had a protective effect.

Mirea and colleagues³ examined the effect of likes on subsequent posting behavior in 3 cross-sectional Twitter datasets. An increase in posting after a like is consistent with a reinforcing effect of likes. Strikingly, the first dataset revealed that the reinforcing effect of likes was increased in individuals who had reported a depression diagnosis. This was conceptually replicated in 2 further samples, where Twitter users had provided depression severity measurements using 2 different standardized self-report scales. Indeed, in one of the datasets, the reinforcing effect was only visible in individuals with heightened depression symptoms. Importantly, these effects seemed specific to depressive symptoms: it was not accounted for by measures of overall mental health, loneliness, social withdrawal, or compulsive/intrusive thoughts.

These findings are noteworthy in a number of ways. First, with regard to depression itself: laboratory studies have long shown a reduction of reinforcement sensitivity in depression, conceptually in keeping with the subjective experience and clinical description of anhedonia.⁸ The fact that, here, reinforcement effects are increased in individuals with depressive symptoms suggests that these effects are not global and vary importantly depending on the setting, or in fact depend on specific goals.⁹ It may also be that patients with depression feel compelled to pursue rewards that are no longer aligned with well-being—a focus of psychotherapies. Interestingly, a recent preprint examining these relationships in a longitudinal dataset found that an increased sensitivity to social media likes predicted declines in mental health, but not vice versa.¹⁰

Second, addictions are clinically characterized by a narrowing of interest on the abused substance.¹¹ In this study,³ there were no separate measures of addiction. However, a broad reduction in interests (seen in reported symptoms of depression) combined with a heightened reward effect in response to social media likes would certainly be the kind of pattern to be expected if social media likes had addictive effects. In that sense, it cannot be excluded that what appears as depression in this study³ may reflect a component of the addictive pattern noted in the US judgments against YouTube and Meta.² Note, however, that a recent preprint did not find the neurocognitive signatures of addiction for social media, which we would expect for more established addictions.¹²

What this work emphasizes, however, is that social media designs or affordances, such as likes, interact with aspects of reinforcement learning, such as how we learn from expected rewards, value expected reward, and select actions.⁶ Critically, with such technologies designed to maximize time spent online, as part of the so-called attention economy, they are inherently motivated to ensure highly rewarding experiences that lead to repeated and ongoing engagement. How such short-term reward exploitation relates to longer-term and more cumulative mental health processes is still largely unknown. But this current work³ highlights that there could be concerning influences at play.

Furthermore, these influences might be further magnified in an age of artificial intelligence (AI), which is also designed to maximize use and promote user return. Therefore, this problem is not going away. Indeed, recent work has found that humans prefer AI that is sycophantic far beyond normal human levels¹³ and that this could also negatively influence the accuracy of AI advice given.¹⁴ A better understanding of how short-term reward optimization interacts with reinforce-

ment learning processes as well as longer-term individual and societal health dynamics is critical.

The work³ has limitations. It only considered Twitter likes, but it seems rather unlikely that none of these effects should generalize to other platforms built around likes. The causal direction is an important caveat in any correlational work, as such analyses cannot determine the direction of the causal arrow. The temporal lag of the effects across the 3 studies does suggest that it was preexisting depressive symptoms that in-

creased like sensitivity. However, it cannot be excluded that the reinforcing effect of likes could instead be the cause of increases in depression symptoms, or that a third factor could have caused both. Any causal effect of social media likes on depressive symptoms should be of major concern to policymakers and may be relevant to legal arguments of harm. Conversely, if social media likes increase engagement in certain vulnerable subgroups, then this raises ethical concerns warranting consideration of safeguarding steps.

ARTICLE INFORMATION

Author Affiliations: Applied Computational Psychiatry Lab, Mental Health Neuroscience Department, Division of Psychiatry, Queen Square Institute of Neurology, University College London, London, United Kingdom (Huys); North London Mental Health NHS Foundation Trust, London, United Kingdom (Huys); Max Planck Centre for Computational Psychiatry and Ageing Research, Department for Imaging Neuroscience, Queen Square Institute of Neurology, University College London, London, United Kingdom (Huys); MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom (Orben).

Corresponding Author: Quentin J. M. Huys, MD, PhD, Division of Psychiatry and Max Planck Centre for Computational Psychiatry and Ageing Research, University College London, 10-12 Russell Square, London WC1B 5EH, United Kingdom (q.huys@ucl.ac.uk).

Published Online: July 8, 2026.
doi:10.1001/jamapsychiatry.2026.1373

Conflict of Interest Disclosures: Dr Huys reported grants from Wellcome Trust, the National Institute of Health Research, Koa Health, and Carigest S.A. and personal fees from Aya Technologies and Alto Neuroscience. Dr Orben reported grants from UK Research and Innovation, Wellcome Trust, Nath Family Trust, Huo Family Foundation, Jacobs Foundation, the UK Department for Science, Innovation and Technology, National Institute of Health, University of Cambridge, Emmanuel College and St John's College of the University of Cambridge, the Prudence Trust and the Livelihood Impact Fund, personal fees from SWFGL, Apple University and various schools, and Psychological Science, as well as a paid reviewer

for a range of grant givers. No other disclosures were reported.

REFERENCES

- Haidt J, Rausch Z. Social media is harming young people at a scale large enough to cause changes at the population level (2026). Accessed June 3, 2026. <https://www.worldhappiness.report/ed/2026/social-media-is-harming-adolescents-at-a-scale-large-enough-to-cause-changes-at-the-population-level>
- The Guardian. Meta and YouTube designed addictive products that harmed young people, jury finds. Accessed June 3, 2026. <https://www.theguardian.com/media/2026/mar/25/jury-verdict-us-first-social-media-addiction-trial-meta-youtube>
- Mirea DM, Chen S, Soto KP, et al. The reinforcement effect of social media likes in depression. *JAMA Psych*. Published online July 8, 2026. doi:10.1001/jamapsychiatry.2026.1430
- Lindström B, Bellander M, Schultner DT, Chang A, Tobler PN, Amodio DM. A computational reward learning account of social media engagement. *Nat Commun*. 2021;12(1):1311. doi:10.1038/s41467-020-19607-x
- Brady WJ, McLoughlin K, Doan TN, Crockett MJ. How social learning amplifies moral outrage expression in online social networks. *Sci Adv*. 2021;7(33):7. doi:10.1126/sciadv.abe5641
- Turner G, Ferguson AM, Katiyar T, Palminteri S, Orben A. Old strategies, new environments: reinforcement learning on social media. *Biol Psychiatry*. 2025;97(10):989-1001. doi:10.1016/j.biopsych.2024.12.012
- Mitchell AJ, McGlinchey JB, Young D, Chelminski I, Zimmerman M. Accuracy of specific symptoms in the diagnosis of major depressive disorder in psychiatric out-patients: data from the MIDAS project. *Psychol Med*. 2009;39(7):1107-1116. doi:10.1017/S0033291708004674
- Halahakoon DC, Kieslich K, O'Driscoll C, Nair A, Lewis G, Roiser JP. Reward-processing behavior in depressed participants relative to healthy volunteers: a systematic review and meta-analysis. *JAMA Psychiatry*. 2020;77(12):1286-1295. doi:10.1001/jamapsychiatry.2020.2139
- Hall AF, Browning M, Huys QJM. The computational structure of consummatory anhedonia. *Trends Cogn Sci*. 2024;28(6):541-553. doi:10.1016/j.tics.2024.01.006
- Davidson BC, Turner G, Gunschera LJ, Orben A. Social media reward sensitivity and depressive symptoms over time. *psyArxiv* (2026). doi:10.31234/osf.io/8e3ar_v1
- Koob GF, Volkow ND. Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry*. 2016;3(8):760-773. doi:10.1016/S2215-0366(16)00104-8
- Gunschera LJ, Mehrhof SZ, Veith C, et al. Does social media use show established neurocognitive signatures of addiction? *psyArxiv* (2026). doi:10.31234/osf.io/2eqs8_v1
- Cheng M, Lee C, Khadpe P, Yu S, Han D, Jurafsky D. Sycophantic AI decreases prosocial intentions and promotes dependence. *Science*. 2026;391(6792):eac8352. doi:10.1126/science.aec8352
- Ibrahim L, Hafner FS, Rocher L. Training language models to be warm can reduce accuracy and increase sycophancy. *Nature*. 2026;652(8112):1159-1165. doi:10.1038/s41586-026-10410-0